

# Construction grammar and the corpus-based analysis of discourses

Groom, Nicholas

DOI:

[10.1075/ijcl.00014.gro](https://doi.org/10.1075/ijcl.00014.gro)

License:

None: All rights reserved

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Groom, N 2019, 'Construction grammar and the corpus-based analysis of discourses: the case of the WAY IN WHICH construction', *International Journal of Corpus Linguistics*, vol. 24, no. 3, pp. 291-323.  
<https://doi.org/10.1075/ijcl.00014.gro>

[Link to publication on Research at Birmingham portal](#)

## **Publisher Rights Statement:**

Checked for eligibility: 22/05/2019

Groom, N. (2019) Construction grammar and the corpus-based analysis of discourses: the case of the WAY IN WHICH construction, *International Journal of Corpus Linguistics*, 24(3), 335–367, <https://doi.org/10.1075/ijcl.00014.gro>

© John Benjamins Publishing Company

## **General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## **Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# **Construction Grammar and the corpus-based analysis of discourses: The case of the WAY IN WHICH construction**

Nicholas Groom

Department of English Language and Linguistics,  
University of Birmingham

Construction grammar (CxG) initially arose as a usage-based alternative to nativist theoretical accounts of language, and remains to this day strongly associated with cognitive linguistic theory and research. In this paper, however, I argue that CxG can be seen as offering an equally viable general framework for socially-oriented linguists whose work focuses on the corpus-based analysis of discourses (CBADs). The paper begins with brief reviews of CxG and CBADs as distinctive research traditions, before going on to identify synergies (both potential and actual) between them. I then offer a more detailed case study example, focusing on a usage-based analysis of a newly identified construction, the WAY IN WHICH construction, as it occurs in corpora representing six different academic discourses. The paper concludes by rebutting some anticipated objections to the approach advocated here, and by proposing a new conceptual model for constructionist approaches to CBADs.

**Keywords:** Construction Grammar, discourse analysis, discourses, phraseology

## **1. Introduction**

Since it first emerged in the 1980s, Construction Grammar has consistently been associated with cognitive linguistics as a general theoretical enterprise (cf. Lakoff, 1987; Fillmore et al., 1988; Croft & Cruse, 2004; Evans & Green, 2006; Hoffmann & Trousdale, 2013) and with cognitively-oriented research on topics such as first and second language acquisition (e.g. Goldberg, 1995, 2006; Ellis, 2003; Gries & Wulff, 2005; Gilquin, 2015; Römer et al., 2015; Kerz & Wiechmann, 2016; Ellis et al., 2016) and language change (e.g. Patten, 2012; Hilpert, 2013; Traugott & Trousdale, 2013; Jing-Schmidt & Peng, 2016; Perek, 2016). At the same time, however, there has also been a steady growth of interest in extending Construction Grammar into the more socially and contextually oriented fields of sociolinguistics, pragmatics and discourse analysis (e.g. Fried & Östman, 2005; Bergs & Diewald, 2009;

Östman & Trousdale, 2013; Aijmer, 2016; Nikiforidou & Fischer, 2016). The aim of this paper is to promote the further extension of Construction Grammar into these realms, by making a practical and theoretical case for its applicability to the now burgeoning subfield of corpus-based research that focuses on the analysis of *discourses*, defined as constellations of meanings and values which simultaneously serve to represent things in the world in particular ways, and to represent the social actors who express these meanings and values as members of particular sociocultural groups (cf. Teubert, 2005, 2010; Baker, 2006; Partington et al., 2013; Baker & McEnery, 2015; Taylor & Marchi, 2018).

In order to establish the terms and the remit of the argument in more detail, the paper begins with brief theoretical overviews of the two research traditions in question, and of previous interfaces between them. I then offer a detailed case study example analysis focusing on a newly identified construction, which I shall refer to as the WAY IN WHICH construction (e.g. *this affected the way in which their food was cooked and served*; from the BNC), in corpora representing six different academic discourses. The paper concludes with a critical discussion of some anticipated objections to the approach advocated here.

## 2. Construction Grammar

As Construction Grammar (conventionally abbreviated to CxG) is not (yet) a mainstream approach within the field of discourse analysis, it will be necessary to begin with a very brief outline of this approach for the benefit of readers who may be unfamiliar with its main tenets.

The first point to note about CxG is that, just like the Chomskyan nativist models of language to which it is in almost all other respects diametrically opposed, it is not simply a descriptive grammar but a fully-fledged attempt to address the fundamental theoretical question ‘what is language?’ CxG’s answer to this question is that each human language is an intricately structured, semantically motivated and constantly evolving inventory of symbolic units, referred to as *constructions*. Each construction is symbolic in that it comprises a pairing of a particular form with a particular meaning. These form-meaning pairs are of greatly varying size and schematicity, ranging from single morphemes such as *dog* to clause-length units such as *you can’t teach an old dog new tricks*, and from fully lexicalized expressions such as *you can’t teach an old dog new tricks* to entirely schematic sequences such as *SUBJ<sub>agent</sub> V OBJ<sub>recipient</sub> OBJ<sub>theme</sub>* (i.e. the English ditransitive construction, of which the proverb *you can’t teach an old dog new tricks* constitutes a very particular exemplar). While

it may be convenient for certain purposes to talk informally about ‘words’ or ‘grammatical structures’ as though they are fundamentally different things, such category distinctions have no ontological status in CxG. On the contrary, ‘words’, ‘grammatical structures’ and everything in between are understood as essentially the same kind of unit, differing only in terms of length, complexity or level of generality. CxG is thus a rigorously monist theory of language, which insists that “knowledge of language is to be modelled as knowledge of constructions, *and nothing else in addition*” (Hilpert, 2014: 22, emphasis added). Furthermore, in answering the related question of how knowledge of a first or subsequent language is acquired, CxG does not postulate any genetically endowed ‘language faculty’ of the kind envisaged by Chomskyan generativism. On the contrary, CxG proposes that language acquisition is motivated by social interaction, and facilitated by general learning mechanisms such as the ability to perceive meaningful patterns, regularities and correspondences in perceptual data and to infer productive generalisations from such observations.

While these basic principles are shared by all construction grammarians, it is also important to note that ‘Construction Grammar’ is not a single unified theoretical approach, but rather a family of approaches and perspectives, each of which has its own distinctive emphases and priorities (cf. Croft & Cruse, 2004; Hoffman & Trousdale, 2013). From the perspective of the current paper, the most important contrast is that between ‘usage-based’ approaches such as Cognitive Construction Grammar (CCxG) (Goldberg, 1996) on the one hand, and more ‘formal’ approaches such as Berkeley Construction Grammar (BCG) (Fillmore & Kay, 1995; Kay, 2002; Fillmore, 2013) and Sign-Based Construction Grammar (SBCG) (Boas & Sag, 2012; Michaelis, 2013) on the other. A helpful summary of the most important differences between these two broad traditions is provided by Boas (2013: 248-49):

Whereas in CCxG the existence of constructions in the grammar [is] thought to be motivated by more general properties of human interaction and cognition, BCG and SBCG do not emphasize the role of motivation. [Instead], most work in BCG and SBCG is aimed at finding maximal generalizations without any redundancy. This means that if a particular expression can be accounted for on the basis of constructions already known to exist, then there is no need to postulate a separate construction. Similarly, BCG and SBCG are not concerned with frequencies of constructions. In contrast, CCxG takes a strong usage-based view of the role of frequency and the status of item-specific instances, leading to the idea that even fully regular

patterns may be stored alongside more abstract schematic constructions when they occur with sufficient frequency[.]

Given their commitment to accounting for observations about natural language usage emerging from cognate fields such as psycholinguistics and corpus linguistics, it is perhaps unsurprising to note that usage-based approaches (and CCxG in particular) have become the default choice among corpus-based construction grammarians in recent years (Yoon & Gries, 2016). Accordingly, the usage-based definition of constructions proposed by Goldberg (2006: 5) will be assumed for the purposes of this paper:

Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency[.]

The final proviso specified in this definition – that constructions can be identified as such purely on the basis of their frequency of occurrence – is of particular utility and relevance to discourse researchers, as will be argued in more detail later in this paper.

### **3. The corpus-based analysis of discourses**

Although the terms ‘discourse’ and ‘discourse analysis’ are notorious for meaning different things to different scholars, it is possible to make a broad distinction between three different (although ultimately interrelated) traditions of discourse research. One of these traditions (let us label it DA1 for convenience) focuses on spoken and written text structure, and is represented by subfields and approaches such as textlinguistics and Rhetorical Structure Theory. The second tradition (let us call it DA2) focuses more widely on the relationship between meaning, context and interaction, and is represented by such fields as pragmatics, conversation analysis, interactional sociolinguistics and the ethnography of communication. In the third tradition, the analytical focus broadens further still, towards the identification and explication of systems of representation that, since the early work of Michel Foucault (1970, 1972), have been referred to across the humanities and social sciences as (countable noun)

*discourses*. It is this third tradition of discourse analysis (henceforth DA3) with which the current paper is primarily concerned.

In this latter tradition, discourses are understood as social practices which simultaneously function in two ways. Firstly, they provide humans as social beings with ways of talking about things in the world that effectively ‘construct’ their knowledge and understandings of these things. That is, they both enable and constrain our ability to talk or think about an entity, process, concept or state of affairs in the world around us in a particular way, or indeed to recognise something as having any ontological status at all (Teubert, 2010). Secondly, by participating in a particular discourse, language users are not only viewing and understanding (an aspect of) the world in a certain way, but also at the same time affiliating themselves to, or claiming membership of, a particular social group (Gee, 2014). That is, in taking up a particular discourse, I am effectively constructing myself (and will be constructed by others) as having a particular professional or sociocultural identity, or as having particular political or religious beliefs, or as occupying a particular stance in relation to an issue of cultural or moral significance in my society, or as having certain aesthetic preferences, and so on.

Crucially, both of these functions of discourse – the entity-constructing and the identity-constructing – are predicated on the same two underlying assumptions: firstly, that discourses must be characterized by symbolic forms that make them *recognizable* to language users (and thus also to discourse analysts); and secondly, that at least some of these symbolic forms must perforce to some degree be *conventionalized*, which is to say, repeated in the same or similar surface form across multiple texts produced by different speakers, writers and signers located in different places and at different points in time. It is the identification of these conventionalized features of discourses, collectively referred to as *phraseologies* (cf. Gledhill, 2000; Hunston, 2002), that forms the central task of the corpus-based analysis of discourses as a distinctive research approach (see e.g. Orpin, 2005; Charles, 2006; Gabrielatos & Baker, 2008; Baker & McEnery, 2015; Potts et al., 2015; Brookes & Harvey, 2016; Jaworska, 2016).

It is now widely recognized that corpus linguistics is ideally suited to the phraseological analysis of discourses as described above. First of all, it is clearly preferable for claims about discourses as large-scale social phenomena to be based on evidence from corpora containing large numbers of texts produced by many different people than it is for such claims to be based on readings of individual texts produced by single authors, no matter how detailed, ingenious or persuasive these individual readings might be (Stubbs, 1997;

Toolan, 1997). Secondly, corpus linguistics provides researchers with a wide (and still expanding) range of methods for identifying frequently occurring and/or statistically significant language features in discourse-specific corpora (Baker, 2006; Taylor & Marchi, 2018; Kopaczyk & Tyrkkö, 2018). The only thing that this emergent field currently lacks at present is any consensus regarding which theoretical models of language might be best suited to the corpus-based analysis of discourses, or even whether such analyses need to be grounded in any explicitly articulated theoretical model of language at all. The aim of the remainder of this paper is to argue that some form of explicit theoretical grounding is indeed desirable, and to make a case for CxG as a particularly apt choice for this purpose.

Before moving on, however, a brief note on nomenclature is in order. The reader may already have noticed that I am using the term ‘corpus-based analysis of discourses’ to refer to an approach that is more commonly known in the research literature either as ‘corpus-based discourse analysis’, or ‘corpus-assisted discourse studies’. To clarify, I prefer to reserve the label ‘corpus based discourse analysis’ for use as an umbrella term encompassing all three of the main traditions of discourse analysis as discussed above. This is in recognition of the fact that corpus-based methods have been applied to subfields as diverse as textlinguistics (e.g. Hoey & O’Donnell, 2008, 2015), pragmatics (e.g. Aijmer, 2015, 2018; Rühlemann, 2018; Rühlemann & Clancy, 2018), genre analysis (e.g. Biber et al., 2007; Upton & Cohen, 2009; Groom & Grieve, 2019), narrative and conversation analysis (e.g. Toolan, 2009; Morton et al., 2011; Koester, 2012; Rühlemann, 2014), as well as to the analysis of discourses in the ‘DA3’ sense described above. In this increasingly heterogeneous context, it is also useful to have access to a more specific term as a means of referring explicitly and exclusively to corpus-based research in the ‘DA3’ tradition, as opposed to other kinds of corpus-based discourse research. For this purpose, I use the label ‘corpus-based analysis of discourses’ (henceforth CBADs) in preference to ‘corpus-assisted discourse studies’ (CADS), as it more accurately describes the status of corpus analysis in this approach, at least as I conceive it. Specifically, it signals a view of corpus analysis not as ‘assisting’ or ‘informing’ more traditional approaches to ‘DA3’, but as the central activity of the whole enterprise.

#### **4. Previous applications of CxG to CBADs**

As mentioned briefly in the introduction to this paper, constructionist approaches have been applied to a growing number of subfields within discourse analysis in recent years. In

particular, CxG frameworks have been successfully applied to studies in information structure and discourse patterning (e.g. Lambrecht, 1994; Michaelis & Lambrecht, 1996; Östman, 2005), context and pragmatics (e.g. Kay, 2004; Fried & Östman, 2005; Bergs & Diewald, 2009; Jing-Schmidt, 2017), dialogue analysis and conversation analysis (e.g. Linell, 2009; Fischer, 2015) and register and genre studies (e.g. Antonopoulou & Nikiforidou, 2011; Nikiforidou & Fischer, 2016; Hoffman & Bergs, 2018). However, all of these studies are contributions to either the ‘DA1’ or ‘DA2’ traditions of discourse analysis as defined earlier. ‘DA3’, in contrast, has thus far been almost entirely neglected in CxG research. In fact, the only published CxG study that addresses this third domain of discourse studies in any direct way is a contrastive corpus-based analysis by Wulff et al. (2007) of the *into*-causative construction (as in *injury forced him into premature retirement* or *the Government was jolted into action*) in British and American English. Although Wulff et al. describe their study as a comparison of national ‘varieties’ rather than ‘discourses’, it is reasonable to regard their work as a ‘DA3’ study insofar as it pursues the idea that systematic differences in patterns of construction usage may be indicative of deep-seated differences in British and American worldviews and identities. As the authors themselves put it:

Bearing the basic assumption in mind that meaning construction is conceptualization ... , which is, among other things, fundamentally shaped by a speaker’s cultural input, an interesting question that has hitherto not been investigated in much detail in Construction Grammar is: to what extent is the meaning potential of a construction variety-specific?  
(Wulff et al., 2007: 278)

Using the methodology of collocation analysis (Stefanowitsch & Gries, 2003; Gries & Stefanowitsch, 2004), Wulff et al. (2007) study variations in the usage of the *into*-causative construction across two newspaper corpora representing British and American English, and find that this construction is significantly more frequently used in their British data to construe causation in terms of (negatively evaluated) physical force or psychological intimidation, whereas this construction is also frequently used in their American data (but not in the British data) to construe causation in terms of (often positively evaluated) verbal persuasion. They also find a marked preference for what they call ‘movement-initializing’ cause predicates (e.g. *sting*, *provoke*, *chivvy*) in British English, and a contrasting preference for ‘movement-restricting’ cause predicates (e.g. *snooper*, *rope*, *frustrate*) in their American corpus. The authors speculate that this latter contrast “may confirm the commonplace



perception that British culture lacks the strong and explicit emphasis on mobility as an essential condition for a happy and free life as we find it in American culture” (Wulff et al. 2007: 279).

Wulff et al. are of course well aware of the limitations of basing such claims on the analysis of individual linguistic features, and are thus careful to conclude by calling for further studies “to see to what extent these differences also show up in other causative constructions” (Wulff et al. 2007: 279). Regrettably, however, this challenge has not yet been taken up by any other CxG researchers, nor have any subsequent studies systematically pursued Wulff et al.’s more general hypothesis about the relationship between constructions and cultural values.

This neglect of CBADs perspectives in the CxG research community is mirrored by a corresponding lack of engagement with CxG on the part of CBADs researchers. Indeed, to my knowledge the only CBADs study to have made explicit use of a constructionist analytical framework thus far is an unpublished PhD thesis by Turo Hiltunen (2010). As with Wulff et al. (2007), the aim of Hiltunen’s study is to investigate whether differences in construction usage across corpora representing different cultural groups can be interpreted as reflecting differences in the ways in which these cultural groups construe (aspects of) the world around them. However, whereas Wulff et al. (2007) draw broad comparisons between speakers of American and British English, Hiltunen’s focus is on the much more specialized discourses of different academic disciplines, as represented by four corpora of research articles in the fields of Law, Literary Criticism, Medicine and Physics respectively.

The part of Hiltunen’s thesis that makes the most thoroughgoing use of a CxG approach centres on a comparative analysis of the *as*-predicative construction (as in *The Egyptians regarded time as a succession of recurring phases* or *The fire is being treated as arson*; both from the BNC) across the abovementioned four corpora. Hiltunen finds significant differences in the usage of this construction across the four fields in question, and argues that these differences can be mapped onto canonical distinctions (originating in the higher education studies literature) between the epistemologically ‘hard’ (i.e. positivist) academic fields of the natural and physical sciences and the epistemologically ‘soft’ (i.e. hermeneutic) fields of the humanities and social sciences (Biglan, 1973a, 1973b). Researchers in ‘hard’ fields tend to work within universally agreed and sharply defined research paradigms, and regard the business of knowledge creation in their discipline as a matter of discovering facts by conducting empirical (and usually experimental) research. In Becher & Trowler’s (2001) highly influential terms, the epistemologies of ‘hard’ fields can

thus be described as “cumulative”, “atomistic” and “concerned with universals, quantities, [and] simplification”. This being the case, Hiltunen argues, it is not surprising to find that the *as*-predicative is predominantly used by physicists and medical researchers “for reporting their own research activities” (p.238), as illustrated in Examples 1 and 2 (taken from Hiltunen’s data) below:

(1) Individual irradiated cells were classified as either clonogenic or nonclonogenic  
(Physics)

(2) We used the midpoint of LVAD enrollment as the dividing point for comparing the 2  
cohorts. (Medicine)

In ‘Soft’ fields, in contrast, there is no single agreed theoretical paradigm for researchers to conform to, and knowledge creation is seen not as a matter of discovering facts or solving problems, but of offering or evaluating different perspectives on perennial questions or canonical texts. In Becher & Trowler’s (2001: 36) terms, the epistemologies of ‘soft’ fields can thus be described as “[r]eiterative”, “holistic” and “concerned with particulars, qualities [and] complication”. These qualities, according to Hiltunen (2010), explain why the *as*-predicative “is much more likely to be used for either advancing a particular claim, or reporting an assertion that has previously been made by someone else”, as illustrated in Examples 3 and 4:

(3) Casebooks generally treat the economic approach as an “exotic perspective” (Law)

(4) We may read it, I propose, either as allegory about the way in which the intensities of  
experience felt as deeply private are also a social gesture, or as aesthetic allegory  
about another kind of publication of private vision. (Literary Criticism)

As Hiltunen (2010) also notes, the underlying semantics of this construction (crudely, ‘*somebody conceptualizes something as something*’) make it a particularly apt linguistic resource for expressing statements of interpretation or evaluation. This in turn explains why the *as*-predicative construction occurs by far the most frequently in Hiltunen’s corpus of academic writing in Literary Criticism, a discipline whose central concern is with the interpretation and evaluation of literary texts.

In summary, the review of previous research above indicates that constructionist approaches have already been successfully applied in empirical studies of discourses at very different levels of socio-cultural specificity. However, it must be conceded that neither of the studies previous discussed above is a full synthesis of constructionist and discourse analytic perspectives. Wulff et al. (2007) make no direct reference to discourse analytic concepts in their study, and Hiltunen (2010) imports aspects of CxG into his discourse research without fully committing to CxG as a theory; indeed, he explicitly describes his overall approach as “theory-neutral” (Hiltunen 2010: 91). In the following section, therefore, it will be necessary to offer a more detailed case study analysis, not only as a means of demonstrating more concretely how CxG can be applied in a CBADs context, but also as an inaugural contribution to what I hope will be a new tradition of research.

## **5. Case study: epistemological variation and the WAY IN WHICH construction**

The case study example to be presented in this section focuses on the analysis of academic discourses. Following essentially the same line of enquiry as Hiltunen (2010), it tests the hypothesis that systematic patterns of variability in construction usage will be observed across corpora representing different academic discourses, and that these patterns of variability can be interpreted as both reflecting and constructing fundamental epistemological differences across these discourses.

The construction chosen for analysis here will be referred to in this paper as the WAY IN WHICH construction. This construction is illustrated by Examples 5 – 7 below, all of which have been taken from the BNC:

- (5) she will have to promote changes in the way in which the markets go about their business.
- (6) The ways in which meaning is expressed must always be central to any language study
- (7) ... the need, for example, to reform the appallingly undemocratic way in which Northern Ireland business is handled at Westminster

This construction has been chosen for present purposes because it has already been identified (albeit in non-constructionist terms) as a characteristic feature of written academic English, notably by Biber et al. (1999). Multi-word sequences based on the trigrams *way in which* and *ways in which* have also been identified as frequently occurring in several studies of academic discourses using a ‘lexical bundles’ methodology (e.g. Biber et al., 2004; Cortes, 2004; Hyland, 2008; Durrant, 2017). However, as this construction has never been previously discussed or even identified as a construction in the previous CxG literature, it will first be necessary to establish that it has constructional status in its own right, and is not merely a ‘construct’ of a more general construction located at a higher level of schematicity.

### 5.1 A closer look at the WAY IN WHICH construction

On the face of it, the morphosyntactic pattern underlying Examples 5 – 7 above would seem to constitute a clear form-meaning pairing, and thus a strong candidate for ‘constructionhood’. In terms of its form, this pattern is a complex noun phrase consisting of three main elements: the head noun *WAY*, followed by the relative adverbial *in which*, followed by a finite clause. The head noun itself is open to all the standard options pertaining to countable nouns, i.e. it may occur in singular or plural form and may be preceded by a definite or indefinite article, or (when plural) by no article at all. The head noun is also often premodified adjectivally, as in Example (7) above; in the BNC, for instance, the node term *way/ways in which* is preceded by at least one adjective in nearly a quarter of all instances.<sup>1</sup> In terms of its meaning and function, this candidate construction serves to focus attention on *how* the process expressed in the finite clause happens or is the case. If the head noun is premodified (as in Example (7) above), it is the premodifying element(s) that provide the ‘manner’ information. If the head noun is not premodified (as in Examples (5) and (6) above), the ‘manner’ meaning is usually left entirely unspecified in the construction itself, and is thus either assumed as already shared knowledge, or predicted (cf. Tadros, 1994) as information that will be provided at some subsequent point in the unfolding text. In some cases, however, the ‘manner’ meaning is stated explicitly in the main clause element of the construction itself. To illustrate this, consider Example (8) below:

- (8) The debate was characterized by widespread dissatisfaction with the way in which the UK had announced Sir David's resignation without appointing a replacement (BNC)

As can be seen, the clause-final prepositional phrase ‘without appointing a replacement’ describes how Sir David’s resignation was announced, and in so doing also explains why there was ‘widespread dissatisfaction’ with the way this resignation was announced.

Taken together, these syntactic and semantic properties can be represented schematically, as shown in Figure 1.

<b>Semantics:</b>	MANNER	PROCESS
<b>Syntax:</b>	[DET] [ADJ] <i>WAY in which</i>	Finite Clause
<b>Examples:</b>	<i>the way in which</i>	<i>the charges are levied</i>
	<i>The ways in which</i>	<i>ideas and beliefs in societies constrain the behaviour of their members</i>
	<i>the appallingly undemocratic way in which</i>	<i>Northern Ireland business is handled at Westminster</i>
	<i>the way in which</i>	<i>the UK had announced Sir David's resignation without appointing a replacement</i>

**Figure 1:** Schematic representation of the WAY IN WHICH construction

One significant shortcoming of the above account, however, is that it fails to acknowledge that ‘*WAY in which* + finite clause’ is just one of three variant ‘manner head noun’ forms featuring *WAY* as head, the other two being ‘*WAY + that* + finite clause’ and ‘*WAY* + finite clause’ (i.e. with no relativizing element). Consider, by way of illustration, Examples 9 – 11 below:

(9) she will have to promote changes in the **way in which** the markets go about their business.

(10) she will have to promote changes in the **way that** the markets go about their business.

(11) she will have to promote changes in the **way** the markets go about their business.

These three variants are syntactically very similar, and there seems to be no discernible change in semantics or in level of formality (in register terms) from one variant to the next. Indeed, it could even be argued that these three variants are better understood as *allostructions*, “variant structural realizations of a construction that is left partially underspecified” (Cappelle, 2006: 18; cf. Perek, 2015) – the MANNER HEAD NOUN construction, perhaps. What, then, is the justification for focusing exclusively on the ‘WAY *in which* + finite clause’ variant as a construction in its own right, as I propose to do in the case study analysis below?

The simplest answer to this question is that ‘WAY *in which* + finite clause’ can be viewed as satisfying the last of Goldberg’s three criteria as discussed earlier: “Any linguistic pattern is recognized as a construction ... *as long as [it] occur[s] with sufficient frequency*” (Goldberg, 2006: 5; emphasis added).

That is, ‘WAY *in which* + finite clause’ qualifies as a construction because its frequency profile – or, more precisely, its frequency *distribution* profile – indicates that English speakers perceive it as having a distinctive status, which in turn means (from a usage-based perspective, at least) that it forms a part of speaker knowledge that must be accounted for on its own terms.

It is important to note that this interpretation goes somewhat beyond Goldberg’s (2006) formulation as quoted above. Goldberg’s concept of “sufficient frequency” is usually interpreted in terms of absolute frequency in the language as a whole, and Goldberg herself makes no explicit claim that a strong association between a pattern and a particular genre in itself constitutes evidence for constructionhood in that genre. It is nevertheless my contention that the criterion of sufficient frequency *can* plausibly and usefully be interpreted from a usage-based perspective as meaning ‘occurring with sufficient frequency *in a particular discourse or genre*’.

Empirical support for this claim comes from Biber et al. (1999: 628), who present a comparative corpus-based analysis of the three variant forms ‘WAY *in which* + finite clause’, ‘WAY + *that* + finite clause’ and ‘WAY + finite clause’. While Biber et al (1999) note that all three of these variant forms are frequently attested in academic prose, they find that ‘WAY *in which* + finite clause’ – but not the other two variant forms – is found to be *exclusively* associated with academic prose as a broad register variety. To check this finding, I also carried out a distributional analysis for the search query *way in which* in the BNC, using the ‘Distribution’ function in BNCweb. The relevant section of this analysis is provided in Appendix 1. Although the BNC analysis paints a somewhat less extreme picture, Biber et al.

(1999)'s essential observation that the WAY IN WHICH construction is strongly associated with academic writing (and dispreferred in news journalism, prose fiction and conversation) holds true. This suggests that it seems to be part of English speakers' knowledge that 'WAY *in which* + finite clause' is strongly primed for use in academic prose, and equally strongly negatively primed for use in these other registers (Hoey, 2005). On the basis of these observations, therefore, I propose that it is valid to regard the morphosyntactic pattern 'WAY *in which* + finite clause' as a construction, and to focus exclusively on this construction in the analysis that follows.

## 5.2 Data and methods

The question to be investigated here follows up on the *LGSWE* account of the WAY IN WHICH construction by asking whether its usage is quantitatively and qualitatively variable across different academic disciplines, and if so, whether such variability can be interpreted as both reflecting and helping to construct different 'ways of knowing' in different fields of study. The methodology of the study is based on concordance analyses of the node terms *way in which* and *ways in which* across six corpora of academic research articles representing six very different disciplinary discourses: Cell Biology and Electrical Engineering, representing the physical and natural sciences; Economics and Sociology, representing social sciences; and English Literature and History, representing the humanities. A summary of the composition of each corpus is provided in Appendix 2.

The two humanities corpora were compiled by the author, and consist of articles published in leading journals between 2010 and 2014. The other four corpora also contain articles from leading journals, but were compiled by another researcher (Oakey, 2008, 2009)<sup>2</sup>, and represent a somewhat older sampling period (from 1998 to 2004). It should also be acknowledged here that the six corpora are of varying size and composition in terms of the number of texts and the number of word tokens they contain, as well as in the number of journals sampled for each discipline. These differences notwithstanding, I would submit that the six corpora remain adequately comparable for the purposes of the current paper.

An initial quantitative analysis was carried out in order to establish the frequency distribution of the WAY IN WHICH construction across the six corpora. Singular and plural forms of the construction were analysed separately. A qualitative concordance analysis of the full set of hits for each query in each corpus was then conducted, following a coding

procedure that will be explained in more detail below. All corpus analyses were carried out using *AntConc* version 3.5.0 (Anthony, 2018).

### 5.3 Quantitative analysis

The results of the initial frequency distribution analysis of the WAY IN WHICH construction across the six corpora are presented in Table 1. Also provided in Table 1 is a set of benchmark figures for general reference purposes. These figures (representing ‘academic prose’ as a broad register category along the lines of the LGWSE analysis discussed earlier) were obtained by adding the results for all six corpora together and dividing the totals by six.

**Table 1:** Frequency distribution of the WAY IN WHICH construction across six discipline-specific corpora

Corpus	Frequency per million words		TOTAL
	<i>way in which</i>	<i>ways in which</i>	
Sociology	56.98	114.75	171.73
English Literature	50.97	69.09	120.06
History	34.49	44.74	79.23
‘Academic prose’	26.86	39.70	66.57
Economics	15.36	8.08	23.44
Cell Biology	1.89	1.11	3.00
Electrical Engineering	1.49	0.44	1.93

Far from being characteristic of ‘academic prose’ in general, the current analysis finds that the WAY IN WHICH construction varies dramatically and systematically across the six disciplinary discourses studied here. The most striking general observation is that the WAY IN WHICH construction occurs very frequently in the humanities, and even more so in the ‘soft’ social science field of sociology, but is actually *uncharacteristic* of academic discourse in hard knowledge fields.

Although the figures for History are only slightly higher than the benchmark frequency, the WAY IN WHICH construction occurs at almost twice the average rate in the discourse of English Literature, and is nearly three times more frequent than the benchmark figure in Sociology. In stark contrast to this, the WAY IN WHICH construction occurs in the discourse of Economics only a third as frequently as might be expected from a general



register perspective, and – perhaps most strikingly of all – it is hardly ever attested in pure science discourses such as Cell Biology and Electrical Engineering.

These patterns of variability are highly consistent with Becher & Trowler’s (2001) characterisations of epistemological variation across academic disciplines as discussed earlier in this article. Given that the principal function of the WAY IN WHICH construction is to focus attention on statements of manner, that is, on *how* things manifest themselves, it is not surprising to find that this construction is strongly preferred in ‘soft’ disciplines, where the creation of new knowledge tends to focus on “particulars, qualities [and] complication.” Becher & Trowler’s account also offers to explain why the WAY IN WHICH construction is hardly ever used in ‘hard’ disciplines. In these discourses, the focus is much more on “universals, quantities, [and] simplification”; that is, researchers are principally interested in discovering new facts about the external world, and not so concerned with nuancing understandings of already established phenomena.

This emphasis in humanities fields on providing detailed, nuanced and complex accounts of disciplinary phenomena (and the corresponding drive towards simplification and generalization in the sciences) also seems to offer a plausible explanation for the other main observation emerging from the analysis presented in Table 1: that the plural variant *ways in which* occurs more frequently than does the singular variant *way in which* in Sociology, English Literature and History, whereas the reverse is the case in Economics, Cell Biology and Electrical Engineering.

#### 5.4 Qualitative analysis

The quantitative analysis reported above does not tell whole story, however. A qualitative inspection of the concordance data reveals a more subtle, but no less systematic, difference in the usage of the WAY IN WHICH construction across the hard/soft epistemological divide.

As explained in Section 5.1, the canonical function of the WAY IN WHICH construction is to comment or focus attention on *how* a phenomenon occurs or is the case. That this phenomenon *does* occur or is the case is, in general, not at issue. To illustrate, consider Examples 12 - 14:

- (12) the way in which Cabinet minutes were recorded, Attlee claimed, stressed the role of the prime minister (History)

(13) the specific organization of welfare state institutions and the way in which policy measures are implemented also affect the structuring of the life course. (Sociology)

(14) *Caleb Williams*'s original publication context was thus crucial to the ways in which the novel was received. (English Literature)

In Examples (12) and (13), it would be decidedly odd to deny that senior government meetings are minuted, or that government policies are (at least sometimes) implemented. Likewise, the novel in question in Example (14) is known to have caused a political stir on its publication in 1794, so there can be no question that it was 'received' in a variety of ways by the reading public of its day.

However, closer analysis of individual concordance lines reveals that there are instances where the statement in the main clause does not have this 'to be taken for granted' status. Consider Examples 15 – 17:

(15) Work of the last decade has explored the ways in which *Mrs. Dalloway*'s prose style constitutes a critique of liberal imperialism's dominant epistemology. (English Literature)

(16) The idea that the Kemalist project of modernization as the bearer of democracy, rights, justice, and truth can provide the vehicle for democratization is called into question by the way in which it positions minorities and rights-asserting groups as particularisms subordinate to its own universalism. (Sociology)

(17) Whereas Eric Evans and many subsequent historians have viewed tithes as illustrative of the way in which the clergy were increasingly distanced and separated from lay society, this article has demonstrated that ... tithe relationships continued to bind the inhabitants of every parish. (History)

In Example (15), the reader does not have to be an expert on Virginia Woolf's fiction (or even need to know that *Mrs Dalloway* is a novel by Virginia Woolf) to recognise that not every literary critic will agree with the claim that the prose style of this novel constitutes a critique of liberal imperialism's dominant epistemology. Similarly, there is a clear

discrepancy in Example (16) between the Kemalist project's decidedly upbeat view of itself and the more critical view of this movement that is being put forward by the writer. The non-factive status of the underlined proposition in Example (17) is even more obvious; the writer here explicitly states that their own research undermines the arguments previously proposed by Eric Evans and other historians.

In these examples, then, the main purpose of the WAY IN WHICH construction seems to be the opposite of the canonical usage as described earlier; that is, the writer seems to be using the construction not to focus the reader's attention on how things manifest themselves, but to argue that these things are actually the case at all. Accordingly, we can call this usage *proposition-focused*, in contrast to the canonical *manner-focused* variant.

Interestingly, it transpires that proposition-focused instances can be identified by means of a simple substitution test. In all cases, the three word sequence *way in which* can be substituted by the two-word sequences *fact that* or *idea that*, without changing the overall meaning of the text, as illustrated in Examples (18)–(20).

(18) Eric Evans and many subsequent historians have viewed tithes as illustrative of the **fact that** the clergy were increasingly distanced and separated from lay society ...

(19) Work of the last decade has explored the **idea that** Mrs. Dalloway's prose style constitutes a critique of liberal imperialism's dominant epistemology.

(20) The idea that the Kemalist project of modernization as the bearer of democracy, rights, justice, and truth can provide the vehicle for democratization is called into question by the **fact that** it positions minorities and rights-asserting groups as particularisms subordinate to its own universalism.

This substitution does not work with canonical 'manner-focused' instances of the WAY IN WHICH construction; in such cases, the reworded statement either makes no sense at all, or entirely changes the original meaning, as can be seen in Examples (21) and (22) below:

(21) \*The **fact/idea that** the axial-flux density changes radially across the active region was calculated by 3-D FEA

(22) \*the **fact/idea that** Cabinet minutes were recorded, Attlee claimed, stressed the role of the prime minister

This substitution test provided the basis for a qualitative investigation into the distribution of these two semantic variants across the six corpora. This investigation centred on an analysis of full sets of concordance results for the search terms *way in which* and *ways in which* in each corpus, and involved the manual inspection of the context of each individual concordance line using *AntConc's File View* tool. In lieu of an inter-rater reliability test, the analysis adopted a stringently conservative policy of coding any instance that could not easily be identified by the substitution test described above as manner-focused by default. The results of this qualitative analysis are presented in Table 2.

**Table 2:** Distribution of manner-focused and proposition-focused WAY IN WHICH construction variants across six discipline-specific corpora

Corpus	Manner-focused			Proposition-focused		
	<i>way in which</i>	<i>ways in which</i>	<i>in TOTAL</i>	<i>way which</i>	<i>in ways in which</i>	<i>in TOTAL</i>
English	250/481	240/651	490/1132	231/481	411/651	642/1132
Literature	(52%)	(36.9%)	(43.3%)	(48%)	(63.1%)	(56.7%)
History	162/275	194/358	356/633	113/275	164/358	277/633
	(58.9%)	(54.2%)	(56.2%)	(41.1%)	(45.8%)	(43.8%)
Sociology	173/289	433/582	606/871	116/289	151/582	267/871
	(59.9%)	(74.1%)	(69.6%)	(40.1%)	(25.9%)	(30.7%)
Economics	76/76	40/40	116/116	0/76	0/40	0/116
	(100%)	(100%)	(100%)	(0%)	(0%)	(0%)
Cell Biology	29/29	17/17	46/46	0/29	0/17	0/46
	(100%)	(100%)	(100%)	(0%)	(0%)	(0%)
Electrical	10/10	3/3	13/13	0/10	0/3	0/13
Engineering	(100%)	(100%)	(100%)	(0%)	(0%)	(0%)

As can be seen, the proposition-focused usage is the exclusive preserve of the discourses of “soft” disciplines; it is not attested in Economics, Cell Biology or Electrical Engineering at all. Among the soft fields, this usage is particularly prominent in English Literature, where it accounts for nearly 57% of all instances. The most likely explanation for this is that this usage is particularly well suited to putting forward propositions that are purely interpretative, and not claims to any kind of objective ‘truth’. The plausibility of interpretative claims in soft fields in general, and in English Literature in particular, rests on the amassing of detailed insights and observations drawn from close readings of texts. This being the case, it is

plausible to argue that the WAY IN WHICH construction is frequently used by literary critics both to put forward an interpretative claim, and at the same time to signal that textual evidence in support of this interpretation will be provided. That is, a statement such as Example (23) simultaneously means both “I will argue that Capote eliminates, or at least reduces, many of the strategic delays ordinarily imposed by the hermeneutic code” and “I will provide examples to show how Capote eliminates, or at least reduces, many of the strategic delays ordinarily imposed by the hermeneutic code”.

(23) I [will] then discuss the way in which Capote eliminates, or at least reduces, many of the strategic delays ordinarily imposed by the hermeneutic code (English Literature)

To an unsympathetic observer, this rhetorical strategy may seem to be little more than a form of argumentative sleight of hand – a way of smuggling in a dubious proposition under the reader’s nose by disguising it as already shared knowledge. Within the context of humanities discourses, however, this particular instance of construction usage is better understood (and will no doubt be understood by members of these discourse communities) as a form of epistemological shorthand (‘I will offer an interpretation and attempt to demonstrate its plausibility through detailed analysis’), and not as an attempt to deceive.

In summary, the case study analysis presented in this section has shown that the WAY IN WHICH construction is strongly associated with the disciplinary discourses of the humanities and ‘soft’ social sciences, and can thus be regarded as a linguistic resource for constructing and transmitting the epistemological values of these discourses. The analysis also finds that, on closer qualitative inspection, this construction is exploited in more than one way in the ‘soft’ fields. In some cases the underlying proposition is agreed or at least uncontroversial, in which case the function of the construction is to introduce or predict the addition of circumstantial details of manner. In other cases, it simultaneously serves to put forward a proposition which is open to dispute, and to signal that evidence in support of this proposition will be provided at a subsequent point in the text.

## **6. Remaining issues and objections**

The aim of the empirical ‘proof of concept’ analysis presented in the previous section was to demonstrate that usage-based construction grammar can be applied rigorously and

productively to the quantitative and qualitative corpus-based analysis of academic and other specialized discourses. Irrespective of the persuasiveness of this case study analysis, however, it is inevitable that misgivings will remain in some quarters about the more general appropriacy of CxG as a theoretical framework for CBADs. Indeed, these misgivings may well explain why constructionist approaches have not been taken up by more CBADs researchers since the publication of the ground-breaking studies by Wulff et al. (2007) and Hiltunen (2010) reviewed earlier. In this section, therefore, I wish to discuss two of the most likely sets of reservations about the synthesis of discourse and constructionist research perspectives advocated in this paper. Each of these will be discussed in turn below.

## 6.1 Terminological issues

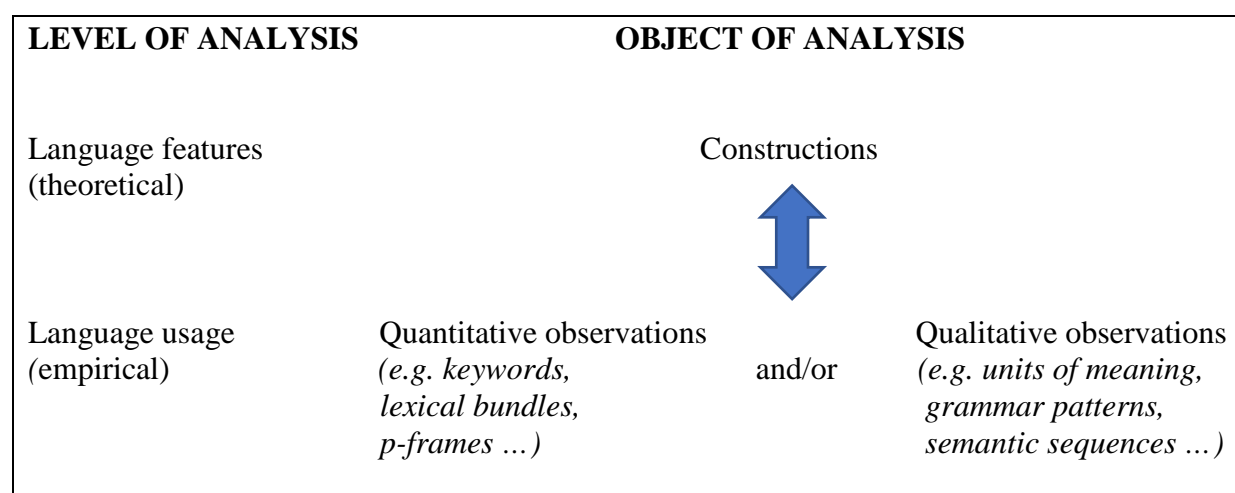
Although a relatively new field of studies, CBADs is already overflowing with specialized terminology (some of it at least partly overlapping) relating to the kinds of linguistic features that researchers in this field typically focus on: collocations, collocation networks, concgrams, formulaic sequences, grammar patterns, idioms, keywords, lexical bundles, lexical items, lexical phrases, n-grams, p(hrase)-frames, phraseological units, semantic sequences, skip-grams and units of meaning, among many others. While this embarrassment of conceptual riches certainly attests to the vigour and eclecticism of CBADs as an emerging field of research, it is also problematic in that it militates against the making of potentially useful comparisons across empirical studies that use different sets of terms. From this perspective, it is easy to understand why some CBADs researchers might be reluctant to engage with CxG. To put it bluntly: the field is terminologically overburdened already, so why complicate matters even further by importing yet another battery of technical terms for what appear to be essentially the same underlying phenomena?

In response to this, I would like to suggest that the reverse is actually the case: that adopting a constructionist approach to CBADs actually *simplifies* the conceptual landscape, and clarifies the relationships among concepts at different levels of analysis. As a first step in this argument, it is necessary to establish that the objection sketched out above rests on a fundamentally misguided assumption: specifically, that there is a broad theoretical equivalence between the corpus-based terms and concepts listed in the previous paragraph (e.g. collocations, grammar patterns, keywords, lexical bundles, p-frames, units of meaning, etc.) on the one hand, and the conceptual apparatus proposed by CxG on the other. This is

very far from the case in reality. On the contrary, whereas the term ‘construction’ is the central concept in a full and coherent theory of language as a symbolic system of usage-based form-meaning pairs, none of the terms listed in the previous paragraph has any formal theoretical status whatsoever; they are all descriptive terms, firmly rooted in the empirical business of corpus analysis. In some cases, of course, this is nothing more than a statement of the obvious. Keywords, lexical bundles, n-grams and p-frames, for example, are clearly not theoretical linguistic concepts in any strict sense at all; rather, they are simply the names of automatic extraction procedures that do double duty as descriptive labels for the kinds of usage data yielded by these procedures. However, this argument also holds true for the less obviously ‘methodological’ of the corpus-based concepts listed above. To take just one example, ‘(extended) units of meaning’ (Sinclair, 1996, 2004) is one of the most important and influential concepts ever to have emerged from the field of corpus linguistics, but it is a concept that has never been formally defined, nor has it ever been placed within a comprehensive theory of language. What is clear from the examples presented in Sinclair’s own work (e.g. his oft-cited analyses of *naked eye*, *true feelings*, *brook* and *place*) is that units of meaning are idiomatic multi-word phrases which are both highly patterned and yet open to substantial internal variation, and whose meanings are a property of the extended unit as a whole and not reducible to any of its constituent parts. Sinclair (1991, 1996, 2004) points out that such units are very frequent (and probably the default) in natural language data, but does not say how far the concept of units of meaning applies to language as an overarching system. This leaves a range of fundamental questions unanswered. If the notion of units of meaning does not in itself constitute a full and comprehensive account of language, then what are its limits? What other kinds of linguistic unit co-exist with units of meaning, and how (if at all) are they interrelated?

In summary, then, corpus linguistic concepts such as units of meaning, collocations, keywords and lexical bundles describe fundamentally important observations about language in use obtained from corpora, but it must be recognized that they do not – either singly or in combination – amount to anything like a formal or comprehensive theory of language of the kind proposed by CxG. They are ways of talking about what language *does*, and not ways of talking about what language *is*. This should not be seen as a problem, however. On the contrary, my argument is that recognising a distinction between constructions as theoretical and explanatory and corpus-based concepts as empirical and descriptive is enormously helpful and liberating. Not only does it provide clearer distinctions between objects of analysis at different levels of analysis, but it also provides a coherent means of articulating

relationships among these. As mentioned above, constructions are abstract theoretical generalisations, and thus cannot be observed directly in corpora. Conversely, corpus-derived features such as collocations, lexical bundles and units of meaning are empirical observations which lack any explicit theoretical status. Each thus dovetails with the other, yielding a two-level analytical model in which a variety of qualitative and quantitative observational methods are used to extract empirical usage data from corpora, which are then reinterpreted theoretically as constructions. This relationship is represented diagrammatically in Figure 2.



**Figure 2:** Levels and objects of analysis in constructionist CBADs

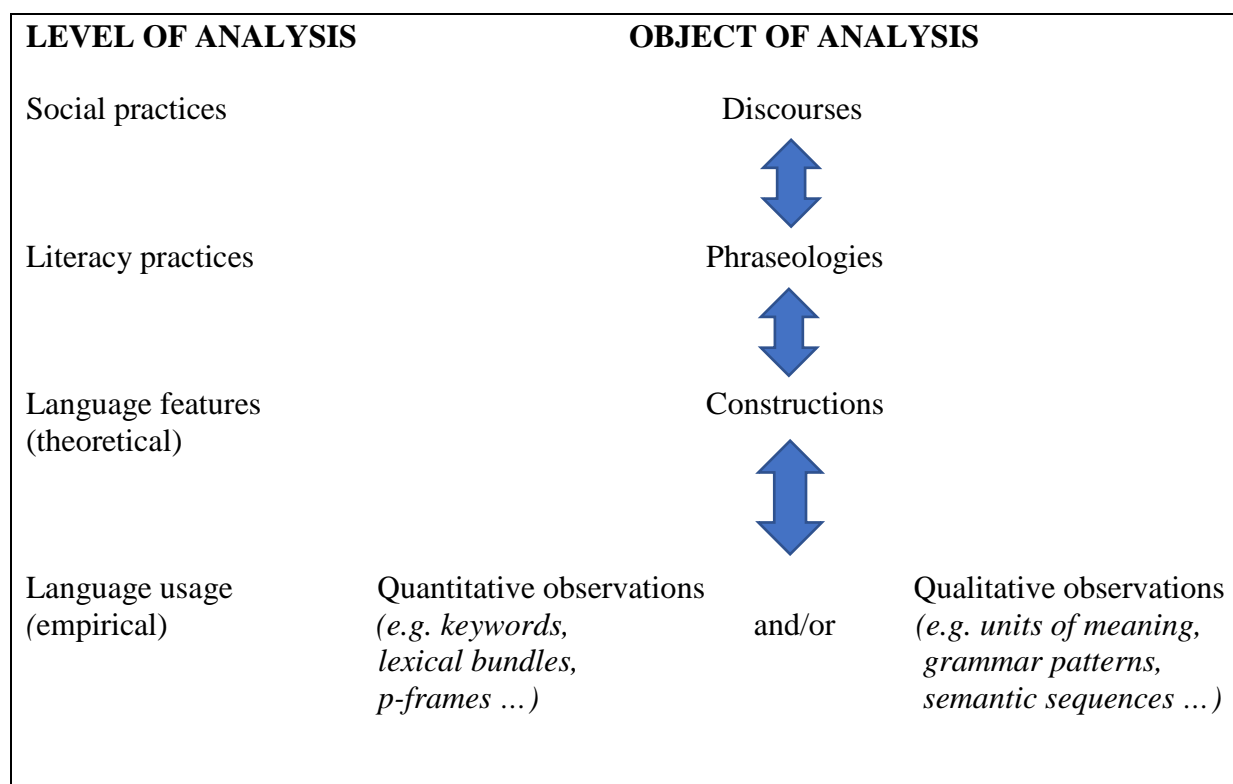
A remaining objection at this point might be that CBADs already has a well-established general term that covers all the features identified by corpus analysis: the concept of *phraseology*, influentially defined by Hunston (2002: 137) as “the tendency of words to occur, not randomly, or even in accordance with grammatical rules only, but in preferred sequences.” (cf. Cowie, 1998; Granger & Paquot, 2008; Gries, 2008). While it is certainly the case that phraseology is the most popular general term for referring to the characteristic linguistic features of discourses, I would argue that there are two fundamental problems with this concept as it is currently used in CBADs, both of which can be solved by replacing it with the concept of constructions.

Firstly, and to recapitulate an earlier argument, ‘phraseology’ turns out on close inspection to be yet another descriptive term, and not a fully-fledged theoretical concept on a par with the concept of ‘construction’. It tells us something important about what language does (i.e. it tends to occur in preferred sequences), but it does not constitute a formal claim



about what language is. As Hunston (2002: 147) puts it, although “phraseology is extremely pervasive ... phraseology alone cannot account for how sentences or utterances are made up.” The concept of construction, in contrast, *does* encompass and account for all kinds of language features, from individual words to complete sentences and utterances, and from fully lexicalized multi-word units to entirely schematic grammatical patterns. In so doing, it renders questions about the boundary between the phraseological and the non-phraseological in language moot, and thus effectively renders the concept of phraseology redundant. (On this point, it is telling that ‘phraseology’ does not even warrant an index entry in Hoffman & Trousdale (2013), the standard reference work on CxG).

Secondly, a review of the CBADs literature finds that phraseology is also frequently used to refer in a much looser and more general way to the whole set of conventionalized language features that have to be learned by anyone who wishes to participate legitimately in any given discourse. This ‘specialized literacy practice’ meaning of the term, defined by Gledhill (2000: 1, 202) as “the preferred way of saying things in a particular discourse” is strongly conterminous with the lay meaning of the term as found in standard dictionary definitions.<sup>3</sup> This conceptualization of phraseology is very useful for CBADs, not only because it captures the idea that discourses have recognisable linguistic characteristics, but also because it recognises that not all contributions to a discourse will be formulaic or even frequently repeated within that discourse (Teubert, 2010). Felicitously, replacing the lexicogrammatical concept of phraseology with the concept of construction means that, instead of being abolished altogether, phraseology can be repositioned at this higher level of theoretical abstraction. This in turn allows us to expand the analytical model proposed in Figure 3 into a full and coherent set of terms and concepts for constructionist approaches to CBADs. This conceptual model is presented in Figure 3.



**Figure 3:** A conceptual model for constructionist approaches to CBADs

Before moving on, it should be noted that the idea of repositioning the concept of phraseology as presented here is not an entirely new one. On the contrary, very similar proposals can be found in Gledhill (2000: 203; see also Gledhill, 2011), and the model presented in Figure 3 is directly inspired by Gledhill's example. The main differences between the two models are that Gledhill's is based around the Hallidayan concept of "lexicogrammatical patterns" rather than constructions, and does not posit an explicit theoretical distinction between phraseology and discourse of the kind envisaged here.

## 6.2 Objections to CxG's 'cognitive commitment'

The second set of objections that I wish to discuss here centres on the CxG's close and abiding relationship with cognitive linguistics. As acknowledged at the beginning of this article, CxG originated as a theoretical alternative to Chomskyan generativism, and shares the generativist aspiration to understand how language is learned by, represented in, and accessed from, the mind of the language user. At issue here, then, is the question of whether and to

what extent this ‘cognitive commitment’ makes CxG fundamentally incompatible with CBADs as a socially-oriented tradition of research.

For some scholars, the answer to this question is a resounding ‘yes’. Teubert (2005: 8), for example, argues not only that “corpus linguistics and cognitive linguistics are two complementary, but ultimately irreconcilable paradigms”, but also that

[t]he mind is no more an object of scientific investigation than the soul ... [and] any attempt to describe what goes on in the mind has to fail because it will run into contradictions that cannot be resolved.

(Teubert, 2010: 43)

For others, objections to CxG’s cognitivist philosophical underpinnings may be motivated by pragmatic rather than ideological concerns. Irrespective of their views on the cogency or otherwise of cognitive theory, researchers who are principally interested in studying how people use language to achieve particular social goals and purposes may understandably see little point or benefit in engaging with a theoretical approach that is avowedly committed to understanding language as a system of knowledge in the minds of individual language users.

The simplest response to both of these objections is to point out that not all construction grammarians regard the ‘cognitive commitment’ as obligatory, or even necessarily relevant in some fields of linguistic enquiry. Here, for example, is Hilpert (2018: 24-25):

While the notion of a schematic pattern strongly invites the idea of a speaker who mentally represents an abstraction over many concrete usage events, it is actually possible for an analyst to remain agnostic about this issue. It can be left open whether abstract patterns have psychological reality or whether they are in fact merely posited as theoretical constructs, much in the way that linguistic notions such as ‘head’, ‘clitic’, or ‘subordinating conjunction’ serve as useful labels for phenomena that capture insights about a linguistic system but that may not correspond to psychologically real categories in the minds of speakers.

Hilpert’s point is made in the context of a review of constructionist approaches to the study of language change, but it applies equally well to the field of discourse analysis. Indeed, the agnostic stance sanctioned by Hilpert is essentially the same as that already adopted by Hiltunen (2010) in his “theory-neutral” study of constructional variability in academic

disciplinary discourses. The position taken by this paper is similar, but not identical to Hiltunen's. Specifically, it does endorse the CxG account of language acquisition, storage and processing, but leaves such matters entirely implicit, on the grounds that they are not directly germane to the research questions that were pursued in Section 5 of this paper.

## **7. Conclusion**

In this paper I have advocated usage-based Construction Grammar as a powerful, flexible and productive framework for carrying out corpus-based research on discourses, defined as ways of saying that are also ways of knowing, being and acting in the social world. The practical case study example at the heart of the paper applied precisely such an approach in an analysis of a newly identified construction, the **WAY IN WHICH** construction, across six corpora representing a range of different academic disciplinary discourses. The analysis found that this construction is very frequently used in humanities and 'soft' social sciences discourses, but hardly used at all in the 'hard' sciences and quantitative social sciences. The analysis also identified two distinct usage variants of the **WAY IN WHICH** construction –manner-focused and proposition-focused – and found that the latter is exclusively associated with the interpretative epistemologies of 'soft' disciplines. Taken together, these observations provide empirical support for the theoretical claim that the epistemologies and phraseologies of academic disciplines are mutually constitutive.

As well as demonstrating the practical viability of applying CxG to CBADs, I have also offered a number of more general theoretical arguments in favour of adopting a constructionist approach to the analysis of discourses. In particular, I have argued that, far from adding unnecessary complexity into the field, basing CBADs around the central theoretical concept of 'construction' allows us to clarify and simplify a number of key terms and concepts, and to bring them into a more coherent overall relationship. Finally, following Hilpert (2018), I have argued that it is perfectly possible for CBADs researchers to embrace usage-based CxG as an overarching theoretical framework without needing to affiliate to, or at least engage explicitly with, its much-vaunted 'cognitive commitment'.

While this paper has generally focused throughout on showing how CBADs can benefit from CxG, I would also like to propose that the reverse is also the case; that is, that CxG stands to benefit substantially from a fuller engagement with the wider perspectives revealed by CBADs. One of the implications of the case study analysis presented in this

paper, for example, is that it indicates that Goldberg's (2006) "sufficient frequency" criterion would benefit from being expanded so that it recognises the observation of a skewed distribution of a feature across discourses as an indicator of constructionhood, as well as (or perhaps even instead of) the frequency of a construction in absolute terms.

Of course, it would be unrealistic as well as presumptuous to argue that all corpus-based discourse analysts should become construction grammarians, or that all construction grammarians should cultivate a keen interest in the analysis of discourses, and it has not been my intention in this paper to propose this. What I am claiming, however, is that CxG is far more compatible with CBADs than it might at first glance seem, and that there are in fact good reasons for regarding CxG as offering the most elegant and compelling theoretical framework for CBADs currently available. If these arguments encourage at least some discourse researchers to consider engaging more fully with CxG than they may have done hitherto, then this paper will have served its main purpose.

## Acknowledgements

I wish to thank the three anonymous reviewers of this paper for their helpful comments and suggestions.

## Notes

1. The search query (*way / ways*) *in which* returns 5279 hits for the whole BNC. Searching for *\_{A}* (*way / ways*) *in which* (i.e. the same query fronted by an adjective) returns 1274 hits, i.e. 24.13% of all instances.

2. I am grateful to David Oakey for generously allowing me to use his data for this study.

3. The *Oxford Online Dictionary*, for example, defines phraseology as "A particular mode of expression, especially one characteristic of a particular speaker or subject area. '*legal phraseology*'."  
<https://en.oxforddictionaries.com/definition/phraseology>

## References

- Aijmer, K. (2015). General extenders in learner language. In N. Groom, M. Charles & S. John (Eds.), *Corpora, Grammar and Discourse: In honour of Susan Hunston* (pp. 211-234). Amsterdam: John Benjamins.
- Aijmer, K. (2016). Pragmatic markers as constructions. The case of *anyway*. In G. Kaltenböck, E. Keizer & A. Lohmann (Eds.), *Outside the Clause: Form and function of extra-clausal constituents*. [Studies in Language Companion Series 178] (pp. 29-58). Amsterdam: John Benjamins.
- Aijmer, K. (2018). 'Corpus pragmatics: From form to function.' In A.H. Jucker, K.P. Schneider & W. Bublitz (Eds.), *Methods in Pragmatics* (pp. 555-586). Berlin: De Gruyter Mouton.
- Anthony, L. (2018). *AntConc* (Version 3.5.0 for Mac OSX). [Computer Software]. <http://www.laurenceanthony.net/software/antconc/>
- Antonopoulou, E., & Nikiforidou, K. (2011). Construction grammar and conventional discourse: A construction-based approach to discursal incongruity. *Journal of Pragmatics*, 43(10), 2594-2609.
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, P., & McEnery, T. (Eds.) (2015). *Corpora and Discourse Studies: Integrating Discourse and Corpora*. Basingstoke: Palgrave Macmillan.
- Becher, T., & Trowler, P. (2001). *Academic Tribes and Territories: Intellectual Enquiry and the Culture of Disciplines*. (2nd ed.). Buckingham: The Society for Research into Higher Education and Open University Press.
- Bergs, A., & Diewald, G. (Eds.) (2009). *Contexts and Constructions*. Amsterdam: John Benjamins.
- Biber, D., Connor, U., & Upton, T.A. (2007). *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: John Benjamins.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at . . . : Lexical bundles in university teaching and textbooks.' *Applied Linguistics*, 25, 371-405.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *The Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.
- Biglan, A. (1973a). The characteristics of subject matter in different scientific areas. *Journal of Applied Psychology*, 57(3), 195-203.

- Biglan, A. (1973b). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology*, 57(3), 204-213.
- Boas, H.C. (2013). Cognitive Construction Grammar. In T. Hoffman & G. Trousdale (Eds.) *The Oxford Handbook of Construction Grammar* (pp. 233-252). Oxford: Oxford University Press.
- Boas, H.C., & Sag, I. (Eds.) (2012). *Sign-Based Construction Grammar*. Stanford: CSLI Publications.
- Brookes, G., & Harvey, K. (2016). Examining the discourse of mental illness in a corpus of online advice-seeking messages. In L. Pickering, E. Friginal, & S. Staples (Eds.) *Talking at Work: Corpus-Based Explorations of Workplace Discourse* (pp. 209-234). Basingstoke: Palgrave Macmillan.
- Cappelle, B. (2006). Particle placement and the case for 'allostructions'. *Constructions*. <https://www.constructions.uni-osnabrueck.de/wp-content/uploads/2014/06/2006-SI-Cappelle22-80-1-PB.pdf>
- Charles, M. (2006). Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes*, 25(3), 310-331.
- Cortes, V. (2004). Lexical bundles in published and student disciplinary writing: Examples from history and biology. *English for Specific Purposes*, 23(4), 397-423.
- Cowie, A.P. (1998). *Phraseology: Theory, Analysis and Applications*. Oxford: Clarendon Press.
- Croft, W., & Cruse, D.A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Durrant, P. (2017). Lexical bundles and disciplinary variation in university students' writing: Mapping the territories. *Applied Linguistics*, 38(2), 165-193.
- Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 63–103). Oxford: Blackwell.
- Ellis, N.C., Römer, U., & O'Donnell, M.B. (2016). *Usage-Based Approaches to Language Acquisition and Processing: Cognitive and Corpus Investigations of Construction Grammar*. [Language Learning Monograph Series]. Malden, MA: Wiley.
- Evans, V., & Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

- Fillmore, C.J. (2013). 'Berkeley Construction Grammar.' In T. Hoffman & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 111-132). Oxford: Oxford University Press.
- Fillmore, C.J., & Kay, P. (1995) *Construction Grammar*. Manuscript, University of California, Berkeley.
- Fillmore, C.J., Kay, P., & O'Connor, M.C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, 64(3), 501-538.
- Fischer, K. (2015). 'Conversation, construction grammar, and cognition.' *Language and Cognition*, 7(4), 563-588.
- Foucault, M. (1970). *The Order of Things*. London: Tavistock.
- Foucault, M. (1972). *The Archaeology of Knowledge*. London: Tavistock.
- Fried, M., & Östman, J.-O. (2005). Construction Grammar and spoken language: The case of pragmatic particles. *Journal of Pragmatics*, 37(11), 1752-1778.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: a corpus analysis of discursive constructions of refugees and asylum seekers in the UK Press 1996-2005. *Journal of English Linguistics* 36(1), 5-38.
- Gee, J.P. (2014). *An Introduction to Discourse Analysis: Theory and Method*. (4<sup>th</sup> ed.). Abingdon: Routledge.
- Gilquin, G. (2015). The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory*, 11(1), 51-88.
- Gledhill, C. (2000). *Collocations in Science Writing*. Tübingen: Gunter Narr.
- Gledhill, C. (2011). 'The 'lexicogrammar' approach to analysing phraseology and collocation in ESP texts.' *Anglais de Spécialité*, 59, 5-23.
- Goldberg, A.E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago and London: University of Chicago Press.
- Goldberg, A.E. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- Goldberg, A.E. (2013). 'Constructionist approaches.' In T. Hoffman & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 15-31). Oxford: Oxford University Press.
- Granger, S., & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Eds.), *Phraseology: An Interdisciplinary Perspective* (pp. 27-50). Amsterdam: John Benjamins.



- Gries, S.T. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Eds.), *Phraseology: An Interdisciplinary Perspective* (pp. 3-25). Amsterdam: John Benjamins.
- Gries, S.T., & Stefanowitsch, A. (2004). Extending collocation analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, 9(1), 97-129.
- Gries, S.T., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting and corpora. *Annual Review of Cognitive Linguistics*, 3, 182-200.
- Groom, N., & Grieve, J. (2019). The evolution of a legal genre: Rhetorical moves in British patent specifications, 1711 to 1860. In T. Fanego & P. Rodríguez-Puente (Eds.), *Corpus-based Research on Variation in English Legal Discourse* (pp. 201–234). Amsterdam: John Benjamins.
- Hilpert, M. (2013). *Constructional Change in English*. Cambridge: Cambridge University Press.
- Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.
- Hilpert, M. (2018). Three open questions in diachronic construction grammar. In E. Coussé, P. Andersson, & J. Olofsson (Eds.), *Grammaticalization meets construction grammar*. [Constructional Approaches to Language 21] (pp. 21-39). Amsterdam: John Benjamins.
- Hiltunen, T. (2010). *Grammar and Disciplinary Culture: A Corpus-Based Study*. Unpublished PhD Thesis, University of Helsinki, Helsinki, Finland.
- Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. Abingdon: Routledge.
- Hoey, M., & O'Donnell, M.B. (2008). Lexicography, grammar, and textual position. *International Journal of Lexicography*, 21(3), 293-309.
- Hoey, M., & O'Donnell, M.B. (2015). Examining associations between lexis and textual position in hard news stories, or *according to a study by ...* In N. Groom, M. Charles & S. John (Eds.), *Corpora, Grammar and Discourse: In Honour of Susan Hunston* (pp. 117-144). Amsterdam: John Benjamins.
- Hoffmann, T. (2017). Construction grammars. In B. Dancygier (Ed.), *The Cambridge Handbook of Cognitive Linguistics* (pp. 310-329). Cambridge: Cambridge University Press.

- Hoffman, T., & Bergs, A. (2018). 'A construction grammar approach to genre.' *CogniTextes* 18. <http://journals.openedition.org/cognitextes/1032> DOI: 10.4000/cognitextes.1032
- Hoffman, T., & Trousdale, G. (Eds.), (2013). *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Hunston, S. (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, 27(1), 4-21.
- Jaworska, S. (2016). Using a corpus-assisted discourse studies (CADS) approach to investigate constructions of identities in media reporting surrounding mega sports events: The case of the London Olympics 2012. In I.R. Lamond & L. Platt (Eds.), *Critical Event Studies. Leisure Studies in a Global Era* (pp. 149-174). Basingstoke: Palgrave Macmillan.
- Jing-Schmidt, Z. (2017). What are they good for? A constructionist account of counterfactuals in ordinary Chinese. *Journal of Pragmatics*, 113(1), 30-52.
- Jing-Schmidt, Z. & X. Peng. (2016). The emergence of disjunction: A history of constructionalization in Chinese. *Cognitive Linguistics*, 27(1), 101-136.
- Kay, P. (2002). An informal sketch of a formal architecture for construction grammar. *Grammars*, 5, 1-19.
- Kay, P. (2004). Pragmatic aspects of grammatical constructions. In L.R. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 675-700). Oxford: Blackwell.
- Kerz, E., & Wiechmann, D. (2016). Second language construction learning: Investigating domain-specific adaptation in advanced L2 production. *Language and Cognition*, 8(4), 533-565.
- Koester, A. (2012). Corpora and workplace discourse. In K. Hyland, M.H. Chau & M. Handford (Eds.), *Corpus Applications in Applied Linguistics: Current Approaches and Future Directions* (pp. 47-64). London: Continuum.
- Kopaczyk, J., & Tyrkkö, J. (Eds.) (2018). *Applications of Pattern-Driven Methods in Corpus Linguistics*. Amsterdam: John Benjamins.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago, IL: University of Chicago Press.
- Lambrecht, K. (1994). *Information Structure and Sentence Form. A Theory of Topic, Focus, and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.

- Linell, P. (2009). Grammatical constructions in dialogue. In A. Bergs & G. Diewald (Eds.), *Contexts and constructions* (pp. 97–110). Amsterdam: John Benjamins.
- Michaelis, L. (2013). Sign-based construction grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 133-152). Oxford: Oxford University Press.
- Michaelis, L., & Lambrecht, K. (1996). Toward a construction-based model of language function: The case of nominal extraposition. *Language*, 72(2), 215-247.
- Morton, T., Walsh, S., & O’Keeffe, A. (2011). Analyzing university spoken interaction: a corpus linguistics/conversation analysis approach. *International Journal of Corpus Linguistics*, 16(3), 326-345.
- Nikiforidou, K., & Fischer, K. (2016). On the interaction of constructions with register and genre. *Constructions and Frames*, 7(2), 137-147.
- Oakey, D. (2008). The form and function of fixed collocational patterns in research articles in different academic disciplines. Unpublished PhD Thesis, University of Leeds, Leeds, UK.
- Oakey, D. (2009). Fixed collocational patterns in isolexical and isotextual versions of a corpus. In P. Baker (Ed.), *Contemporary Corpus Linguistics* (pp. 140-58). London: Bloomsbury.
- Orpin, D. (2005). Corpus Linguistics and Critical Discourse Analysis: Examining the ideology of sleaze. *International Journal of Corpus Linguistics*, 10(1), 37-61.
- Östman, J.-O. (2005). Construction discourse: A prolegomenon. In J.-O. Östman & M. Fried (Eds.), *Construction Grammars: Cognitive Grounding and Theoretical Extensions* (pp. 121–44). Amsterdam: John Benjamins.
- Östman, J.-O., & Trousdale, G. (2013). Dialects, discourse, and Construction Grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 476-490). Oxford: Oxford University Press.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam: John Benjamins.
- Patten, A. (2012). *The English It-cleft: A Constructional Account and a Diachronic Investigation*. [Topics in English Linguistics 79]. Berlin: De Gruyter Mouton.
- Perek, F. (2015). *Argument Structure in Usage-Based Construction Grammar: Experimental and Corpus-Based Perspectives*. Amsterdam: John Benjamins.

- Perek, F. (2016). Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics*, 54(1), 149–188.
- Potts, A., Bednarek, M., & Caple, H. (2015). How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse & Communication*, 9(2), 149-172.
- Römer, U., O'Donnell, M.B., & Ellis, N.C. (2015). 'Using COBUILD grammar patterns for a large-scale analysis of verb-argument constructions: Exploring corpus data and speaker knowledge.' In N. Groom, M. Charles & S. John (eds.). *Corpora, Grammar and Discourse: In Honour of Susan Hunston* (pp. 43-71). Amsterdam: John Benjamins.
- Rühlemann, C. (2014). *Narrative in English Conversation: A Corpus Analysis of Storytelling*. Cambridge: Cambridge University Press.
- Rühlemann, C. (2018). *Corpus Linguistics for Pragmatics: A Guide for Research*. Abingdon: Routledge.
- Rühlemann, C., & Clancy, B. (2018). Corpus linguistics and pragmatics. In C. Ilie & N.R. Norrick (Eds.), *Pragmatics and its Interfaces* (pp. 241-266). Amsterdam: John Benjamins.
- Sinclair, J.M. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J.M. (1996). The search for units of meaning. *Textus*, 9(1): 75–106.
- Sinclair, J.M, (2004) *Trust the Text: Language, Corpus and Discourse*. R. Carter (Ed.). London: Routledge.
- Stefanowitsch, A., & Gries, S.T. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8(2), 209-43.
- Stubbs, M. (1997). Whorf's children: Critical comments on critical discourse analysis (CDA). In A. Ryan & A. Wray (Eds.), *Evolving models of language* (pp. 100-116). Clevedon: Multilingual Matters.
- Tadros, A. (1994). Predictive categories in expository text. In M. Coulthard (Ed.), *Advances in written text analysis* (pp. 69-82). London: Routledge.
- Taylor, C. & Marchi, A. (Eds.) (2018). *Corpus Approaches to Discourse: A Critical Review*. Abingdon: Routledge.
- Teubert, W. (2005). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1-13.

- Teubert, W. (2010). *Meaning, Discourse and Society*. Cambridge: Cambridge University Press.
- Toolan, M. (1997). What is critical discourse analysis and why are people saying such terrible things about it? *Language and Literature*, 6(2), 83-103.
- Toolan, M. (2009). *Narrative Progression in the Short Story: A Corpus Stylistic Approach*. Amsterdam: John Benjamins.
- Traugott, E.C., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.
- Upton, T.A., & Cohen, M.A. (2009). An approach to corpus-based discourse analysis: The move analysis as example. *Discourse Studies*, 11(5), 585–605.
- Wulff, S. (2008). *Rethinking Idiomaticity: A Usage-based Approach*. London: Continuum.
- Wulff, S., Stefanowitsch, A., & Gries, S.T. (2007). Brutal Brits and persuasive Americans: variety-specific meaning construction in the *into*-causative. In G. Radden, K. Köpcke, T. Berg & P. Siemund (Eds.), *Aspects of Meaning Construction* (pp. 265-81). Amsterdam: John Benjamins.
- Yoon, J. & Gries, S.T. (2016). *Corpus-based Approaches to Construction Grammar*. [Constructional Approaches to Language 19]. Amsterdam: John Benjamins.

**Appendix 1: BNCweb ‘Distribution’ analysis results for the query *way in which* (‘derived text type’ data only).**

Derived text type:				
Category	No. of words	No. of hits	Dispersion (over files)	Frequency per million words
Academic prose	15,778,028	<a href="#">1,292</a>	321/497	81.89
Other spoken material	6,175,896	<a href="#">277</a>	127/755	44.85
Non-academic prose and biography	24,178,674	<a href="#">952</a>	324/744	39.37
Other published written	17,924,109	<a href="#">545</a>	179/710	30.41

material				
Unpublished written material	4,466,673	<a href="#">131</a>	63/251	29.33
Newspapers	9,412,174	<a href="#">123</a>	78/486	13.07
Fiction and verse	16,143,913	<a href="#">103</a>	68/452	6.38
Spoken conversation	4,233,962	<a href="#">15</a>	10/153	3.54
<b>total</b>	<b>98,313,429</b>	<b>3,438</b>	<b>1,170/4,048</b>	<b>34.97</b>